SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: (US only) ANTALIS Toni Marie and HOOPER John David

(ii) TITLE OF INVENTION: NOVEL MOLECULES

(iii) NUMBER OF SEQUENCES: 30

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
(B) STREET: 400 GARDEN CITY PLAZA
(C) CITY: GARDEN CITY
(D) STATE: NEW YORK
(E) COUNTRY: USA
(F) ZIP: 11530

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: 09/023,942
(B) FILING DATE: 13-FEB-1998
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: PO5101/97
(B) FILING DATE: 13-FEB-1997

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: PP0422/97
(B) FILING DATE: 18-NOV-1997

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: International PCT Application
(B) FILING DATE: 13-FEB-1998

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: DIGIGLIO, FRANK S
(B) REGISTRATION NO: 31,346
(C) REFERENCE/DOCKET NUMBER: 11168

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (516) 742 4343
(B) TELEFAX: (516) 742 4366
(C) TELEX: 230 901 SANS UR

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:
(A) NAME/KEY: modified base
(B) LOCATION: 15
(D) OTHER INFORMATION: N equals Inosine

(ix) FEATURE:
(A) NAME/KEY: modified base
(B) LOCATION: 18
(D) OTHER INFORMATION: N equals Inosine

(ix) FEATURE:
(A) NAME/KEY: modified base
(B) LOCATION: 21
(D) OTHER INFORMATION: N equals Inosine

(ix) FEATURE:
(A) NAME/KEY: modified base
(B) LOCATION: 24
(D) OTHER INFORMATION: N equals Inosine

(ix) FEATURE:
(A) NAME/KEY: modified base
(B) LOCATION: 27
(D) OTHER INFORMATION: N equals Inosine

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ACAGAATTCT GGGTNGTNAC NGCNGCNCAY TG 32

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:
(A) NAME/KEY: modified base
(B) LOCATION: 12
(D) OTHER INFORMATION: N equals Inosine

(ix) FEATURE:
(A) NAME/KEY: modified base
(B) LOCATION: 15
(D) OTHER INFORMATION: N equals Inosine

(ix) FEATURE:
(A) NAME/KEY: modified base
(B) LOCATION: 18
(D) OTHER INFORMATION: N equals Inosine

(ix) FEATURE:
(A) NAME/KEY: modified base
(B) LOCATION: 21
(D) OTHER INFORMATION: N equals Inosine

(ix) FEATURE:
(A) NAME/KEY: modified base
(B) LOCATION: 27
(D) OTHER INFORMATION: N equals Inosine

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

ACAGAATTCA RNGGNCCNCC NSWRTCNCC 29

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1094 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 17..955

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
CGCGGGAGAG GAGGCC ATG GGC GCG CGC GGG GCG CTG CTG CTG GCG CTG          49
                  Met Gly Ala Arg Gly Ala Leu Leu Leu Ala Leu
                    1               5                  10

CTG CTG GCT CGG GCT GGA CTC AGG AAG CCG GAG TCG CAG GAG GCG GCG        97
Leu Leu Ala Arg Ala Gly Leu Arg Lys Pro Glu Ser Gln Glu Ala Ala
            15                  20                  25

CCG TTA TCA GGA CCA TGC GGC CGA CGG GTC ATC ACG TCG CGC ATC GTG       145
Pro Leu Ser Gly Pro Cys Gly Arg Arg Val Ile Thr Ser Arg Ile Val
        30                  35                  40

GGT GGA GAG GAC GCC GAA CTC GGG CGT TGG CCG TGG CAG GGG AGC CTG       193
Gly Gly Glu Asp Ala Glu Leu Gly Arg Trp Pro Trp Gln Gly Ser Leu
    45                  50                  55

CGC CTG TGG GAT TCC CAC GTA TGC GGA GTG AGC CTG CTC AGC CAC CGC       241
Arg Leu Trp Asp Ser His Val Cys Gly Val Ser Leu Leu Ser His Arg
60                  65                  70                  75

TGG GCA CTC ACG GCG GCG CAC TGC TTT GAA ACT GAC CTT AGT GAT CCC       289
Trp Ala Leu Thr Ala Ala His Cys Phe Glu Thr Asp Leu Ser Asp Pro
                80                  85                  90
```

```
TCC GGG TGG ATG GTC CAG TTT GGC CAG CTG ACT TCC ATG CCA TCC TTC      337
Ser Gly Trp Met Val Gln Phe Gly Gln Leu Thr Ser Met Pro Ser Phe
            95                  100                 105

TGG AGC CTG CAG GCC TAC TAC ACC CGT TAC TTC GTA TCG AAT ATC TAT      385
Trp Ser Leu Gln Ala Tyr Tyr Thr Arg Tyr Phe Val Ser Asn Ile Tyr
        110                 115                 120
```

```
CTG AGC CCT CGC TAC CTG GGG AAT TCA CCC TAT GAC ATT GCC TTG GTG       433
Leu Ser Pro Arg Tyr Leu Gly Asn Ser Pro Tyr Asp Ile Ala Leu Val
    125                 130                 135

AAG CTG TCT GCA CCT GTC ACC TAC ACT AAA CAC ATC CAG CCC ATC TGT       481
Lys Leu Ser Ala Pro Val Thr Tyr Thr Lys His Ile Gln Pro Ile Cys
140                 145                 150                 155

CTC CAG GCC TCC ACA TTT GAG TTT GAG AAC CGG ACA GAC TGC TGG GTG       529
Leu Gln Ala Ser Thr Phe Glu Phe Glu Asn Arg Thr Asp Cys Trp Val
                160                 165                 170

ACT GGC TGG GGG TAC ATC AAA GAG GAT GAG GCA CTG CCA TCT CCC CAC       577
Thr Gly Trp Gly Tyr Ile Lys Glu Asp Glu Ala Leu Pro Ser Pro His
            175                 180                 185

ACC CTC CAG GAA GTT CAG GTC GCC ATC ATA AAC AAC TCT ATG TGC AAC       625
Thr Leu Gln Glu Val Gln Val Ala Ile Ile Asn Asn Ser Met Cys Asn
        190                 195                 200

CAC CTC TTC CTC AAG TAC AGT TTC CGC AAG GAC ATC TTT GGA GAC ATG       673
His Leu Phe Leu Lys Tyr Ser Phe Arg Lys Asp Ile Phe Gly Asp Met
    205                 210                 215

GTT TGT GCT GGC AAT GCC CAA GGC GGG AAG GAT GCC TGC TTC GGT GAC       721
Val Cys Ala Gly Asn Ala Gln Gly Gly Lys Asp Ala Cys Phe Gly Asp
220                 225                 230                 235

TCA GGT GGA CCC TTG GCC TGT AAC AAG GAT GGA CTG TGG TAT CAG ATT       769
Ser Gly Gly Pro Leu Ala Cys Asn Lys Asp Gly Leu Trp Tyr Gln Ile
                240                 245                 250

GGA GTC GTG AGC TGG GGA GTG GGC TGT GGT CGG CCC AAT CGG CCC GGT       817
Gly Val Val Ser Trp Gly Val Gly Cys Gly Arg Pro Asn Arg Pro Gly
            255                 260                 265

GTC TAC ACC AAT ATC AGC CAC CAC TTT GAG TGG ATC CAG AAG CTG ATG       865
Val Tyr Thr Asn Ile Ser His His Phe Glu Trp Ile Gln Lys Leu Met
        270                 275                 280
```

```
GCC CAG AGT GGC ATG TCC CAG CCA GAC CCC TCC TGG CCG CTA CTC TTT      913
Ala Gln Ser Gly Met Ser Gln Pro Asp Pro Ser Trp Pro Leu Leu Phe
    285                 290                 295

TTC CCT CTT CTC TGG GCT CTC CCA CTC CTG GGG CCG GTC TGA            955
Phe Pro Leu Leu Trp Ala Leu Pro Leu Leu Gly Pro Val  *
300                 305                 310

GCCTACCTGA GCCCATGCAG CCTGGGGCCA CTGCCAAGTC AGGCCCTGGT TCTCTTCTGT   1015

CTTGTTTGGT AATAAACACA TTCCAGTTGA TGCCTTGCAG GGCATTTTTC AAAAAAAAAA   1075

AAAAAAAAAA AAAAAAAAA                                               1094
```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 312 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
Met Gly Ala Arg Gly Ala Leu Leu Leu Ala Leu Leu Leu Ala Arg Ala
 1               5                  10                  15

Gly Leu Arg Lys Pro Glu Ser Gln Glu Ala Ala Pro Leu Ser Gly Pro
            20                  25                  30

Cys Gly Arg Arg Val Ile Thr Ser Arg Ile Val Gly Gly Glu Asp Ala
        35                  40                  45

Glu Leu Gly Arg Trp Pro Trp Gln Gly Ser Leu Arg Leu Trp Asp Ser
    50                  55                  60

His Val Cys Gly Val Ser Leu Leu Ser His Arg Trp Ala Leu Thr Ala
65                  70                  75                  80
```

```
Ala His Cys Phe Glu Thr Asp Leu Ser Asp Pro Ser Gly Trp Met Val
                 85                  90                  95

Gln Phe Gly Gln Leu Thr Ser Met Pro Ser Phe Trp Ser Leu Gln Ala
            100                 105                 110

Tyr Tyr Thr Arg Tyr Phe Val Ser Asn Ile Tyr Leu Ser Pro Arg Tyr
        115                 120                 125

Leu Gly Asn Ser Pro Tyr Asp Ile Ala Leu Val Lys Leu Ser Ala Pro
    130                 135                 140

Val Thr Tyr Thr Lys His Ile Gln Pro Ile Cys Leu Gln Ala Ser Thr
145                 150                 155                 160

Phe Glu Phe Glu Asn Arg Thr Asp Cys Trp Val Thr Gly Trp Gly Tyr
                165                 170                 175

Ile Lys Glu Asp Glu Ala Leu Pro Ser Pro His Thr Leu Gln Glu Val
            180                 185                 190

Gln Val Ala Ile Ile Asn Asn Ser Met Cys Asn His Leu Phe Leu Lys
        195                 200                 205

Tyr Ser Phe Arg Lys Asp Ile Phe Gly Asp Met Val Cys Ala Gly Asn
    210                 215                 220

Ala Gln Gly Gly Lys Asp Ala Cys Phe Gly Asp Ser Gly Gly Pro Leu
225                 230                 235                 240

Ala Cys Asn Lys Asp Gly Leu Trp Tyr Gln Ile Gly Val Val Ser Trp
                245                 250                 255

Gly Val Gly Cys Gly Arg Pro Asn Arg Pro Gly Val Tyr Thr Asn Ile
            260                 265                 270

Ser His His Phe Glu Trp Ile Gln Lys Leu Met Ala Gln Ser Gly Met
        275                 280                 285

Ser Gln Pro Asp Pro Ser Trp Pro Leu Leu Phe Phe Pro Leu Leu Trp
    290                 295                 300
```

```
Ala Leu Pro Leu Leu Gly Pro Val  *
305                 310
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1100 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 17..961

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
CGCGGGAGAG GAGGCC ATG GGC GCG CGC GGG GCG CTG CTG CTG GCG CTG          49
                  Met Gly Ala Arg Gly Ala Leu Leu Leu Ala Leu
                    1               5                  10

CTG CTG GCT CGG GCT GGA CTC AGG AAG CCG GAG TCG CAG GAG GCG GCG        97
Leu Leu Ala Arg Ala Gly Leu Arg Lys Pro Glu Ser Gln Glu Ala Ala
            15                  20                  25

CCG TTA TCA GGA CCA TGC GGC CGA CGG GTC ATC ACG TCG CGC ATC GTG       145
Pro Leu Ser Gly Pro Cys Gly Arg Arg Val Ile Thr Ser Arg Ile Val
        30                  35                  40

GGT GGA GAG GAC GCC GAA CTC GGG CGT TGG CCG TGG CAG GGG AGC CTG       193
Gly Gly Glu Asp Ala Glu Leu Gly Arg Trp Pro Trp Gln Gly Ser Leu
    45                  50                  55

CGC CTG TGG GAT TCC CAC GTA TGC GGA GTG AGC CTG CTC AGC CAC CGC       241
Arg Leu Trp Asp Ser His Val Cys Gly Val Ser Leu Leu Ser His Arg
60                  65                  70                  75

TGG GCA CTC ACG GCG GCG CAC TGC TTT GAA ACC TAT AGT GAC CTT AGT       289
```

```
Trp Ala Leu Thr Ala Ala His Cys Phe Glu Thr Tyr Ser Asp Leu Ser
                  80                  85                  90

GAT CCC TCC GGG TGG ATG GTC CAG TTT GGC CAG CTG ACT TCC ATG CCA       337
Asp Pro Ser Gly Trp Met Val Gln Phe Gly Gln Leu Thr Ser Met Pro
             95                 100                 105

TCC TTC TGG AGC CTG CAG GCC TAC TAC ACC CGT TAC TTC GTA TCG AAT       385
Ser Phe Trp Ser Leu Gln Ala Tyr Tyr Thr Arg Tyr Phe Val Ser Asn
        110                 115                 120

ATC TAT CTG AGC CCT CGC TAC CTG GGG AAT TCA CCC TAT GAC ATT GCC       433
Ile Tyr Leu Ser Pro Arg Tyr Leu Gly Asn Ser Pro Tyr Asp Ile Ala
    125                 130                 135

TTG GTG AAG CTG TCT GCA CCT GTC ACC TAC ACT AAA CAC ATC CAG CCC       481
Leu Val Lys Leu Ser Ala Pro Val Thr Tyr Thr Lys His Ile Gln Pro
140                 145                 150                 155

ATC TGT CTC CAG GCC TCC ACA TTT GAG TTT GAG AAC CGG ACA GAC TGC       529
Ile Cys Leu Gln Ala Ser Thr Phe Glu Phe Glu Asn Arg Thr Asp Cys
                160                 165                 170

TGG GTG ACT GGC TGG GGG TAC ATC AAA GAG GAT GAG GCA CTG CCA TCT       577
Trp Val Thr Gly Trp Gly Tyr Ile Lys Glu Asp Glu Ala Leu Pro Ser
            175                 180                 185

CCC CAC ACC CTC CAG GAA GTT CAG GTC GCC ATC ATA AAC AAC TCT ATG       625
Pro His Thr Leu Gln Glu Val Gln Val Ala Ile Ile Asn Asn Ser Met
        190                 195                 200

TGC AAC CAC CTC TTC CTC AAG TAC AGT TTC CGC AAG GAC ATC TTT GGA       673
Cys Asn His Leu Phe Leu Lys Tyr Ser Phe Arg Lys Asp Ile Phe Gly
    205                 210                 215

GAC ATG GTT TGT GCT GGC AAT GCC CAA GGC GGG AAG GAT GCC TGC TTC       721
Asp Met Val Cys Ala Gly Asn Ala Gln Gly Gly Lys Asp Ala Cys Phe
220                 225                 230                 235
```

```
GGT GAC TCA GGT GGA CCC TTG GCC TGT AAC AAG GAT GGA CTG TGG TAT       769
Gly Asp Ser Gly Gly Pro Leu Ala Cys Asn Lys Asp Gly Leu Trp Tyr
                240                 245                 250

CAG ATT GGA GTC GTG AGC TGG GGA GTG GGC TGT GGT CGG CCC AAT CGG       817
Gln Ile Gly Val Val Ser Trp Gly Val Gly Cys Gly Arg Pro Asn Arg
            255                 260                 265

CCC GGT GTC TAC ACC AAT ATC AGC CAC CAC TTT GAG TGG ATC CAG AAG       865
Pro Gly Val Tyr Thr Asn Ile Ser His His Phe Glu Trp Ile Gln Lys
        270                 275                 280

CTG ATG GCC CAG AGT GGC ATG TCC CAG CCA GAC CCC TCC TGG CCG CTA       913
Leu Met Ala Gln Ser Gly Met Ser Gln Pro Asp Pro Ser Trp Pro Leu
    285                 290                 295

CTC TTT TTC CCT CTT CTC TGG GCT CTC CCA CTC CTG GGG CCG GTC TGAGCCTACC 968
Leu Phe Phe Pro Leu Leu Trp Ala Leu Pro Leu Leu Gly Pro Val  *
300                 305                 310                 315

TGAGCCCATG CAGCCTGGGG CCACTGCCAA GTCAGGCCCT GGTTCTCTTC TGTCTTGTTT    1028

GGTAATAAAC ACATTCCAGT TGATGCCTTG CAGGGCATTT TTCAAAAAAA AAAAAAAAAA    1088

AAAAAAAAAA AA                                                        1100
```

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 314 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```
Met Gly Ala Arg Gly Ala Leu Leu Leu Ala Leu Leu Leu Ala Arg Ala
  1               5                  10                  15
```

```
Gly Leu Arg Lys Pro Glu Ser Gln Glu Ala Ala Pro Leu Ser Gly Pro
            20                  25                  30

Cys Gly Arg Arg Val Ile Thr Ser Arg Ile Val Gly Gly Glu Asp Ala
        35                  40                  45

Glu Leu Gly Arg Trp Pro Trp Gln Gly Ser Leu Arg Leu Trp Asp Ser
    50                  55                  60

His Val Cys Gly Val Ser Leu Leu Ser His Arg Trp Ala Leu Thr Ala
65                  70                  75                  80

Ala His Cys Phe Glu Thr Tyr Ser Asp Leu Ser Asp Pro Ser Gly Trp
                85                  90                  95

Met Val Gln Phe Gly Gln Leu Thr Ser Met Pro Ser Phe Trp Ser Leu
            100                 105                 110

Gln Ala Tyr Tyr Thr Arg Tyr Phe Val Ser Asn Ile Tyr Leu Ser Pro
        115                 120                 125

Arg Tyr Leu Gly Asn Ser Pro Tyr Asp Ile Ala Leu Val Lys Leu Ser
    130                 135                 140

Ala Pro Val Thr Tyr Thr Lys His Ile Gln Pro Ile Cys Leu Gln Ala
145                 150                 155                 160

Ser Thr Phe Glu Phe Glu Asn Arg Thr Asp Cys Trp Val Thr Gly Trp
                165                 170                 175

Gly Tyr Ile Lys Glu Asp Glu Ala Leu Pro Ser Pro His Thr Leu Gln
            180                 185                 190

Glu Val Gln Val Ala Ile Ile Asn Asn Ser Met Cys Asn His Leu Phe
        195                 200                 205

Leu Lys Tyr Ser Phe Arg Lys Asp Ile Phe Gly Asp Met Val Cys Ala
    210                 215                 220

Gly Asn Ala Gln Gly Gly Lys Asp Ala Cys Phe Gly Asp Ser Gly Gly
225                 230                 235                 240
```

```
Pro Leu Ala Cys Asn Lys Asp Gly Leu Trp Tyr Gln Ile Gly Val Val
                245                 250                 255

Ser Trp Gly Val Gly Cys Gly Arg Pro Asn Arg Pro Gly Val Tyr Thr
            260                 265                 270

Asn Ile Ser His His Phe Glu Trp Ile Gln Lys Leu Met Ala Gln Ser
        275                 280                 285

Gly Met Ser Gln Pro Asp Pro Ser Trp Pro Leu Leu Phe Phe Pro Leu
    290                 295                 300

Leu Trp Ala Leu Pro Leu Leu Gly Pro Val
305                 310
```

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 799 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 24..799

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

```
AGTTCAGATG AATGGGACTG TGA GAA CCA TCT GTG ACC AAA TTG ATA CAG          50
                          Glu Pro Ser Val Thr Lys Leu Ile Gln
                            1               5

GAA CAG GAG AAA GAG CCG CGG TGG CTG ACA TTA CAC TCC AAC TGG GAG        98
Glu Gln Glu Lys Glu Pro Arg Trp Leu Thr Leu His Ser Asn Trp Glu
 10                  15                  20                  25

AGC CTC AAT GGG ACC ACT TTA CAT GAA CTT GTA GTA AAT GGG CAG TCT       146
```

```
Ser Leu Asn Gly Thr Thr Leu His Glu Leu Val Val Asn Gly Gln Ser
                30                  35                  40

TGT GAG AGC AGA AGT AAA ATT TCT CTT CTG TGT ACT AAA CAA GAC TGT      194
Cys Glu Ser Arg Ser Lys Ile Ser Leu Leu Cys Thr Lys Gln Asp Cys
            45                  50                  55

GGG CGC CGC CCT GCT GCC CGA ATG AAC AAA AGG ATC CTT GGA GGT CGG      242
Gly Arg Arg Pro Ala Ala Arg Met Asn Lys Arg Ile Leu Gly Gly Arg
        60                  65                  70

ACG AGT CGC CCT GGA AGG TGG CCA TGG CAG TGT TCT CTG CAG AGT GAA      290
Thr Ser Arg Pro Gly Arg Trp Pro Trp Gln Cys Ser Leu Gln Ser Glu
    75                  80                  85

CCC AGT GGA CAT ATC TGT GGC TGT GTC CTC ATT GCC AAG AAG TGG GTT      338
Pro Ser Gly His Ile Cys Gly Cys Val Leu Ile Ala Lys Lys Trp Val
90                  95                  100                 105

GTG ACA GTT GCC CAC TGC TTC GAG GGG AGA GAG AAT GCT GCA GTT TGG      386
Val Thr Val Ala His Cys Phe Glu Gly Arg Glu Asn Ala Ala Val Trp
                110                 115                 120

AAA GTG GTG CTT GGC ATC AAC AAT CTA GAC CAT CCA TCA GTG TTC ATG      434
Lys Val Val Leu Gly Ile Asn Asn Leu Asp His Pro Ser Val Phe Met
            125                 130                 135

CAG ACA CGC TTT GTG AGG ACC ATC ATC CTG CAT CCC CGC TAC AGT CGA      482
Gln Thr Arg Phe Val Arg Thr Ile Ile Leu His Pro Arg Tyr Ser Arg
        140                 145                 150

GCA GTG GTG GAC TAT GAC ATC AGC ATC GTT GAG CTG AGT GAA GAC ATC      530
Ala Val Val Asp Tyr Asp Ile Ser Ile Val Glu Leu Ser Glu Asp Ile
    155                 160                 165

AGT GAG ACT GGC TAC GTC CGG CCT GTC TGC TTG CCC AAC CCG GAG CAG      578
Ser Glu Thr Gly Tyr Val Arg Pro Val Cys Leu Pro Asn Pro Glu Gln
170                 175                 180                 185
```

```
TGG CTA GAG CCT GAC ACG TAC TGC TAT ATC ACA GGC TGG GGC CAC ATG       626
Trp Leu Glu Pro Asp Thr Tyr Cys Tyr Ile Thr Gly Trp Gly His Met
                190                 195                 200

GGC AAT AAA ATG CCA TTT AAG CTG CAA GAG GGA GAG GTC CGC ATT ATT       674
Gly Asn Lys Met Pro Phe Lys Leu Gln Glu Gly Glu Val Arg Ile Ile
            205                 210                 215

TCT CTG GAA CAT TGT CAG TCC TAC TTT GAC ATG AAG ACC ATC ACC ACT       722
Ser Leu Glu His Cys Gln Ser Tyr Phe Asp Met Lys Thr Ile Thr Thr
        220                 225                 230

CGG ATG ATA TGT GCT GGC TAT GAG TCT GGC ACA GTT GAT TCA TGC ATG       770
Arg Met Ile Cys Ala Gly Tyr Glu Ser Gly Thr Val Asp Ser Cys Met
    235                 240                 245

GGT GAC TGG GGC GGT CCG TTG AAT TCT GT                               799
Gly Asp Trp Gly Gly Pro Leu Asn Ser
250                 255
```

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 258 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```
Glu Pro Ser Val Thr Lys Leu Ile Gln Glu Gln Glu Lys Glu Pro Arg
  1               5                  10                  15

Trp Leu Thr Leu His Ser Asn Trp Glu Ser Leu Asn Gly Thr Thr Leu
            20                  25                  30

His Glu Leu Val Val Asn Gly Gln Ser Cys Glu Ser Arg Ser Lys Ile
        35                  40                  45
```

```
Ser Leu Leu Cys Thr Lys Gln Asp Cys Gly Arg Arg Pro Ala Ala Arg
     50                  55                  60

Met Asn Lys Arg Ile Leu Gly Gly Arg Thr Ser Arg Pro Gly Arg Trp
 65                  70                  75                  80

Pro Trp Gln Cys Ser Leu Gln Ser Glu Pro Ser Gly His Ile Cys Gly
                 85                  90                  95

Cys Val Leu Ile Ala Lys Lys Trp Val Val Thr Val Ala His Cys Phe
            100                 105                 110

Glu Gly Arg Glu Asn Ala Ala Val Trp Lys Val Val Leu Gly Ile Asn
        115                 120                 125

Asn Leu Asp His Pro Ser Val Phe Met Gln Thr Arg Phe Val Arg Thr
    130                 135                 140

Ile Ile Leu His Pro Arg Tyr Ser Arg Ala Val Val Asp Tyr Asp Ile
145                 150                 155                 160

Ser Ile Val Glu Leu Ser Glu Asp Ile Ser Glu Thr Gly Tyr Val Arg
                165                 170                 175

Pro Val Cys Leu Pro Asn Pro Glu Gln Trp Leu Glu Pro Asp Thr Tyr
            180                 185                 190

Cys Tyr Ile Thr Gly Trp Gly His Met Gly Asn Lys Met Pro Phe Lys
        195                 200                 205

Leu Gln Glu Gly Glu Val Arg Ile Ile Ser Leu Glu His Cys Gln Ser
    210                 215                 220

Tyr Phe Asp Met Lys Thr Ile Thr Thr Arg Met Ile Cys Ala Gly Tyr
225                 230                 235                 240

Glu Ser Gly Thr Val Asp Ser Cys Met Gly Asp Trp Gly Gly Pro Leu
                245                 250                 255

Asn Ser
```

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2241 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 166..1773

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

```
ATTTAATACG ACTCACTATA GGGAATTTGG CCCTCGAGGA AGAATTCGGC ACGAGGCTGC       60

GGCGCACTGT GAGGGAGTCG CTGTGATCCG GGGCCCCGAA CCCGACTGGA GCTGAAGCGC      120

AGGCTGCGGG GCGCGGAGTC GGGAGGCCTG AGTGTTCCTT CCAGC ATG TCG GAG          174
                                                    Met Ser Glu
                                                      1

GGG GAG TCC CAG ACA GTA CTT AGC AGT GGC TCA GAC CCA AAG GTA GAA         222
Gly Glu Ser Gln Thr Val Leu Ser Ser Gly Ser Asp Pro Lys Val Glu
      5                  10                  15

TCT TCA TCT TCA GCT CCT GGC CTG ACA TCA GTG TCA CCT CCT GTG ACC         270
Ser Ser Ser Ser Ala Pro Gly Leu Thr Ser Val Ser Pro Pro Val Thr
 20                  25                  30                  35

TCC ACA ACC TCA GCT GCT TCC CCA GAG GAA GAA GAA GAA AGT GAA GAT         318
Ser Thr Thr Ser Ala Ala Ser Pro Glu Glu Glu Glu Glu Ser Glu Asp
                 40                  45                  50
```

```
GAG TCT GAG ATT TTG GAA GAG TCG CCC TGT GGG CGC TGG CAG AAG AGG      366
Glu Ser Glu Ile Leu Glu Glu Ser Pro Cys Gly Arg Trp Gln Lys Arg
             55                  60                  65

CGA GAA GAG GTG AAT CAA CGG AAT GTA CCA GGT ATT GAC AGT GCA TAC      414
Arg Glu Glu Val Asn Gln Arg Asn Val Pro Gly Ile Asp Ser Ala Tyr
         70                  75                  80

CTG GCC ATG GAT ACA GAG GAA GGT GTA GAG GTT GTG TGG AAT GAG GTA      462
Leu Ala Met Asp Thr Glu Glu Gly Val Glu Val Val Trp Asn Glu Val
     85                  90                  95

CAG TTC TCT GAA CGC AAG AAC TAC AAG CTG CAG GAG GAA AAG GTT TGT      510
Gln Phe Ser Glu Arg Lys Asn Tyr Lys Leu Gln Glu Glu Lys Val Cys
100                 105                 110                 115

GCT GTG TTT GAT AAT TTG ATT CAA TTG GAG CAT CTT AAC ATT GTT AAG      558
Ala Val Phe Asp Asn Leu Ile Gln Leu Glu His Leu Asn Ile Val Lys
                120                 125                 130

TTT CAC AAA TAT TGG GCT GAC ATT AAA GAG AAC AAG GCC AGG GTC ATT      606
Phe His Lys Tyr Trp Ala Asp Ile Lys Glu Asn Lys Ala Arg Val Ile
            135                 140                 145

TTT ATC ACA GGA TAC ATG TCA TCT GGG AGT CTG AAG CAA TTT CTG AAG      654
Phe Ile Thr Gly Tyr Met Ser Ser Gly Ser Leu Lys Gln Phe Leu Lys
        150                 155                 160

AAG ACC CAA AAG AAC CAC CAG ACG ATG AAT GAA AAG GCA TGG AAG CGT      702
Lys Thr Gln Lys Asn His Gln Thr Met Asn Glu Lys Ala Trp Lys Arg
    165                 170                 175

TGG TGC ACA CAA ATC CTC TCT GCC CTA AGC TAC CTG CAC TCC TGT GAC      750
Trp Cys Thr Gln Ile Leu Ser Ala Leu Ser Tyr Leu His Ser Cys Asp
180                 185                 190                 195

CCC CCC ATC ATC CAT GGG AAC CTG ACC TGT GAC ACC ATC TTC ATC CAG      798
Pro Pro Ile Ile His Gly Asn Leu Thr Cys Asp Thr Ile Phe Ile Gln
                200                 205                 210
```

```
CAC AAC GGA CTC ATC AAG ATT GGC TCT GTG GCT CCT GAC ACT ATC AAC        846
His Asn Gly Leu Ile Lys Ile Gly Ser Val Ala Pro Asp Thr Ile Asn
            215                 220                 225

AAT CAT GTG AAG ACT TGT CGA GAA GAG CAG AAG AAT CTA CAC TTC TTT        894
Asn His Val Lys Thr Cys Arg Glu Glu Gln Lys Asn Leu His Phe Phe
        230                 235                 240

GCA CCA GAG TAT GGA GAA GTC ACT AAT GTG ACA ACA GCA GTG GAC ATC        942
Ala Pro Glu Tyr Gly Glu Val Thr Asn Val Thr Thr Ala Val Asp Ile
    245                 250                 255

TAC TCC TTT GGC ATG TGT GCA CTG GGG ATG GCA GTG CTG GAG ATT CAG        990
Tyr Ser Phe Gly Met Cys Ala Leu Gly Met Ala Val Leu Glu Ile Gln
260                 265                 270                 275

GGC AAT GGA GAG TCC TCA TAT GTG CCA CAG GAA GCC ATC AGC AGT GCC       1038
Gly Asn Gly Glu Ser Ser Tyr Val Pro Gln Glu Ala Ile Ser Ser Ala
                280                 285                 290

ATC CAG CTT CTA GAA GAC CCA TTA CAG AGG GAG TTC ATT CAA AAG TGC       1086
Ile Gln Leu Leu Glu Asp Pro Leu Gln Arg Glu Phe Ile Gln Lys Cys
            295                 300                 305

CTG CAG TCT GAG CCT GCT CGC AGA CCA ACA GCC AGA GAA CTT CTG TTC       1134
Leu Gln Ser Glu Pro Ala Arg Arg Pro Thr Ala Arg Glu Leu Leu Phe
        310                 315                 320

CAC CCA GCA TTG TTT GAA GTG CCC TCG CTC AAA CTC CTT GCG GCC CAC       1182
His Pro Ala Leu Phe Glu Val Pro Ser Leu Lys Leu Leu Ala Ala His
    325                 330                 335

TGC ATT GTG GGA CAC CAA CAC ATG ATC CCA GAG AAC GCT CTA GAG GAG       1230
Cys Ile Val Gly His Gln His Met Ile Pro Glu Asn Ala Leu Glu Glu
340                 345                 350                 355

ATC ACC AAA AAC ATG GAT ACT AGT GCC GTA CTG GCT GAA ATC CCT GCA       1278
Ile Thr Lys Asn Met Asp Thr Ser Ala Val Leu Ala Glu Ile Pro Ala
                360                 365                 370
```

```
GGA CCA GGA AGA GAA CCA GTT CAG ACT TTG TAC TCT CAG TCA CCA GCT      1326
Gly Pro Gly Arg Glu Pro Val Gln Thr Leu Tyr Ser Gln Ser Pro Ala
            375                 380                 385

CTG GAA TTA GAT AAA TTC CTT GAA GAT GTC AGG AAT GGG ATC TAT CCT      1374
Leu Glu Leu Asp Lys Phe Leu Glu Asp Val Arg Asn Gly Ile Tyr Pro
        390                 395                 400

CTG ACA GCC TTT GGG CTG CCT CGG CCC CAG CAG CCA CAG CAG GAG GAG      1422
Leu Thr Ala Phe Gly Leu Pro Arg Pro Gln Gln Pro Gln Gln Glu Glu
    405                 410                 415

GTG ACA TCA CCT GTC GTG CCC CCC TCT GTC AAG ACT CCG ACA CCT GAA      1470
Val Thr Ser Pro Val Val Pro Pro Ser Val Lys Thr Pro Thr Pro Glu
420                 425                 430                 435

CCA GCT GAG GTG GAG ACT CGC AAG GTG GTG CTG ATG CAG TGC AAC ATT      1518
Pro Ala Glu Val Glu Thr Arg Lys Val Val Leu Met Gln Cys Asn Ile
                440                 445                 450

GAG TCG GTG GAG GAG GGA GTC AAA CAC CAC CTG ACA CTT CTG CTG AAG      1566
Glu Ser Val Glu Glu Gly Val Lys His His Leu Thr Leu Leu Leu Lys
            455                 460                 465

TTG GAG GAC AAA CTG AAC CGG CAC CTG AGC TGT GAC CTG ATG CCA AAT      1614
Leu Glu Asp Lys Leu Asn Arg His Leu Ser Cys Asp Leu Met Pro Asn
        470                 475                 480

GAG AAT ATC CCC GAG TTG GCG GCT GAG CTG GTG CAG CTG GGC TTC ATT      1662
Glu Asn Ile Pro Glu Leu Ala Ala Glu Leu Val Gln Leu Gly Phe Ile
    485                 490                 495

AGT GAG GCT GAC CAG AGC CGG TTG ACT TCT CTG CTA GAA GAG ACC TTG      1710
Ser Glu Ala Asp Gln Ser Arg Leu Thr Ser Leu Leu Glu Glu Thr Leu
500                 505                 510                 515

AAC AAG TTC AAT TTT GCC AGG AAC AGT ACC CTC AAC TCA GCC GCT GTC      1758
Asn Lys Phe Asn Phe Ala Arg Asn Ser Thr Leu Asn Ser Ala Ala Val
                520                 525                 530
```

```
ACC GTC TCC TCT TAGAGCTCAC TCGGGCCAGG CCCTGATCTG CGCTGTGGCT          1810
Thr Val Ser Ser
            535

GTCCCTGGAC GTGCTGCAGC CCTCCTGTCC CTTCCCCCCA GTCAGTATTA CCCTGTGAAG    1870

CCCCTTCCCT CCTTTATTAT TCAGGAGGGC TGGGGGGGCT CCCTGGTTCT GAGCATCATC    1930

CTTTCCCCTC CCCTCTCTTC CTCCCCTCTG CACTTTGTTT ACTTGTTTTG CACAGACGTG    1990

GGCCTGGGCC TTCTCAGCAG CCGCCTTCTA GTTGGGGGCT AGTCGCTGAT CTGCCGGCTC    2050

CCGCCCAGCC TGTGTGGAAA GGAGGCCCAC GGGCACTAGG GGAGCCGAAT TCTACAATCC    2110

CGCTGGGGCG GCCGGGGCGG GAGAGAAAGG TGGTGCTGCA GTGGTGGCCC TGGGGGGCCA    2170

TTCGATTCGC CTCAGTTGCT GCTGTAATAA AAGTCTACTT TTTGCTAAAA AAAAAAAAAA    2230

AAAAAAAAAA A                                                         2241
```

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 535 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```
Met Ser Glu Gly Glu Ser Gln Thr Val Leu Ser Ser Gly Ser Asp Pro
 1               5                  10                  15

Lys Val Glu Ser Ser Ser Ser Ala Pro Gly Leu Thr Ser Val Ser Pro
            20                  25                  30

Pro Val Thr Ser Thr Thr Ser Ala Ala Ser Pro Glu Glu Glu Glu Glu
        35                  40                  45
```

```
Ser Glu Asp Glu Ser Glu Ile Leu Glu Glu Ser Pro Cys Gly Arg Trp
     50                  55                  60

Gln Lys Arg Arg Glu Glu Val Asn Gln Arg Asn Val Pro Gly Ile Asp
 65                  70                  75                  80

Ser Ala Tyr Leu Ala Met Asp Thr Glu Glu Gly Val Glu Val Val Trp
                 85                  90                  95

Asn Glu Val Gln Phe Ser Glu Arg Lys Asn Tyr Lys Leu Gln Glu Glu
            100                 105                 110

Lys Val Cys Ala Val Phe Asp Asn Leu Ile Gln Leu Glu His Leu Asn
        115                 120                 125

Ile Val Lys Phe His Lys Tyr Trp Ala Asp Ile Lys Glu Asn Lys Ala
    130                 135                 140

Arg Val Ile Phe Ile Thr Gly Tyr Met Ser Ser Gly Ser Leu Lys Gln
145                 150                 155                 160

Phe Leu Lys Lys Thr Gln Lys Asn His Gln Thr Met Asn Glu Lys Ala
                165                 170                 175

Trp Lys Arg Trp Cys Thr Gln Ile Leu Ser Ala Leu Ser Tyr Leu His
            180                 185                 190

Ser Cys Asp Pro Pro Ile Ile His Gly Asn Leu Thr Cys Asp Thr Ile
        195                 200                 205

Phe Ile Gln His Asn Gly Leu Ile Lys Ile Gly Ser Val Ala Pro Asp
    210                 215                 220

Thr Ile Asn Asn His Val Lys Thr Cys Arg Glu Glu Gln Lys Asn Leu
225                 230                 235                 240

His Phe Phe Ala Pro Glu Tyr Gly Glu Val Thr Asn Val Thr Thr Ala
                245                 250                 255

Val Asp Ile Tyr Ser Phe Gly Met Cys Ala Leu Gly Met Ala Val Leu
            260                 265                 270
```

Glu Ile Gln Gly Asn Gly Glu Ser Ser Tyr Val Pro Gln Glu Ala Ile
275 280 285

Ser Ser Ala Ile Gln Leu Leu Glu Asp Pro Leu Gln Arg Glu Phe Ile
290 295 300

Gln Lys Cys Leu Gln Ser Glu Pro Ala Arg Arg Pro Thr Ala Arg Glu
305 310 315 320

Leu Leu Phe His Pro Ala Leu Phe Glu Val Pro Ser Leu Lys Leu Leu
325 330 335

Ala Ala His Cys Ile Val Gly His Gln His Met Ile Pro Glu Asn Ala
340 345 350

Leu Glu Glu Ile Thr Lys Asn Met Asp Thr Ser Ala Val Leu Ala Glu
355 360 365

Ile Pro Ala Gly Pro Gly Arg Glu Pro Val Gln Thr Leu Tyr Ser Gln
370 375 380

Ser Pro Ala Leu Glu Leu Asp Lys Phe Leu Glu Asp Val Arg Asn Gly
385 390 395 400

Ile Tyr Pro Leu Thr Ala Phe Gly Leu Pro Arg Pro Gln Gln Pro Gln
405 410 415

Gln Glu Glu Val Thr Ser Pro Val Val Pro Pro Ser Val Lys Thr Pro
420 425 430

Thr Pro Glu Pro Ala Glu Val Glu Thr Arg Lys Val Val Leu Met Gln
435 440 445

Cys Asn Ile Glu Ser Val Glu Glu Gly Val Lys His His Leu Thr Leu
450 455 460

Leu Leu Lys Leu Glu Asp Lys Leu Asn Arg His Leu Ser Cys Asp Leu
465 470 475 480

Met Pro Asn Glu Asn Ile Pro Glu Leu Ala Ala Glu Leu Val Gln Leu
485 490 495

Gly Phe Ile Ser Glu Ala Asp Gln Ser Arg Leu Thr Ser Leu Leu Glu
500 505 510

Glu Thr Leu Asn Lys Phe Asn Phe Ala Arg Asn Ser Thr Leu Asn Ser
515 520 525

Ala Ala Val Thr Val Ser Ser
530 535

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GCACAGTCGA CCAAGCCGGA GTCGCAGAG 29

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GCACAAAGCT TGCCAGGAGG GGTCTGGCTG 30

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GCACAACCAT GGCCAAGCCG GAGTCGCAGG AG 32

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GCACAAGATC TCCAGGAGGG GTCTGGCTG 29

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Lys Pro Glu Ser Gln Glu Ala Ala Pro Leu Ser Gly Pro Cys
5 10

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Glu Asp Ala Glu Leu Gly Arg Trp Pro Trp Gln Gly Ser Leu Arg Leu Trp Asp
5 10 15
Cys

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Gly Tyr Ile Lys Glu Asp Glu Ala Leu Pro Ser Pro His Thr Leu Gln Cys
5 10 15

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

GCACAGGTAC CGAGGCCATG GGCGCGCGC 29

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 50 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GCACATCTAG ATCAGTGGTG GTGGTGGTGG TGGACCGGCC CCAGGAGTGG 50

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GCACAGCGGC CGCGAGGCCA TGGGCGCGCG C 31

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 52 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GCACAGCGGC CGCTCAGTGG TGGTGGTGGT GGTGCCAGGA GGGGTCTGGC TG 52

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

CTGACTTCCA TGCCATCCTT 20

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GCTCACGACT CCAATCTGAT 20

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Arg Ile Val Gly Gly
5

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 959 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 2..856

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

```
C GAC CTA TTG TCA GGG CCC TGC GGT CAC AGG ACC ATC CCT TCC CGT          46
  Asp Leu Leu Ser Gly Pro Cys Gly His Arg Thr Ile Pro Ser Arg
    1               5                  10                  15

ATA GTG GGT GGC GAT GAT GCT GAG CTT GGC CGC TGG CCG TGG CAA GGG        94
Ile Val Gly Gly Asp Asp Ala Glu Leu Gly Arg Trp Pro Trp Gln Gly
                20                  25                  30

AGC CTG CGT GTA TGG GGC AAC CAC TTA TGT GGC GCA ACC TTG CTC AAC       142
Ser Leu Arg Val Trp Gly Asn His Leu Cys Gly Ala Thr Leu Leu Asn
            35                  40                  45

CGC CGC TGG GTG CTT ACA GCT GCC CAC TGC TTC CAA AAG GAT AAC GAT       190
Arg Arg Trp Val Leu Thr Ala Ala His Cys Phe Gln Lys Asp Asn Asp
        50                  55                  60
```

```
CCT TTT GAC TGG ACA GTC CAG TTT GGT GAG CTG ACT TCC AGG CCA TCT      238
Pro Phe Asp Trp Thr Val Gln Phe Gly Glu Leu Thr Ser Arg Pro Ser
     65                  70                  75

CTC TGG AAC CTA CAG GCC TAT TCC AAC CGT TAC CAA ATA GAA GAT ATT      286
Leu Trp Asn Leu Gln Ala Tyr Ser Asn Arg Tyr Gln Ile Glu Asp Ile
 80                  85                  90                  95

TTC CTG AGC CCC AAG TAC TCG GAG CAG TAT CCC AAT GAC ATA GCC CTG      334
Phe Leu Ser Pro Lys Tyr Ser Glu Gln Tyr Pro Asn Asp Ile Ala Leu
                100                 105                 110

CTG AAG CTG TCA TCT CCA GTC ACC TAC AAT AAC TTC ATC CAG CCC ATC      382
Leu Lys Leu Ser Ser Pro Val Thr Tyr Asn Asn Phe Ile Gln Pro Ile
            115                 120                 125

TGC CTC CTG AAC TCC ACG TAC AAG TTT GAG AAC CGA ACT GAC TGC TGG      430
Cys Leu Leu Asn Ser Thr Tyr Lys Phe Glu Asn Arg Thr Asp Cys Trp
        130                 135                 140

GTG ACC GGC TGG GGG GCT ATT GGA GAA GAT GAG AGT CTG CCA TCT CCC      478
Val Thr Gly Trp Gly Ala Ile Gly Glu Asp Glu Ser Leu Pro Ser Pro
    145                 150                 155

AAC ACT CTC CAG GAA GTG CAG GTA GCT ATT ATC AAC AAC AGC ATG TGT      526
Asn Thr Leu Gln Glu Val Gln Val Ala Ile Ile Asn Asn Ser Met Cys
160                 165                 170                 175

AAC CAT ATG TAC AAA AAG CCA GAC TTC CGC ACG AAC ATC TGG GGA GAC      574
Asn His Met Tyr Lys Lys Pro Asp Phe Arg Thr Asn Ile Trp Gly Asp
                180                 185                 190

ATG GTT TGC GCT GGC ACT CCT GAA GGT GGC AAG GAT GCC TGC TTT GGT      622
Met Val Cys Ala Gly Thr Pro Glu Gly Gly Lys Asp Ala Cys Phe Gly
            195                 200                 205

GAC TCG GGA GGA CCC TTG GCC TGC GAC CAG GAT ACG GTG TGG TAT CAG      670
Asp Ser Gly Gly Pro Leu Ala Cys Asp Gln Asp Thr Val Trp Tyr Gln
        210                 215                 220
```

```
GTT GGA GTT GTG AGC TGG GGA ATA GGC TGT GGT CGC CCC AAT CGC CCT       718
Val Gly Val Val Ser Trp Gly Ile Gly Cys Gly Arg Pro Asn Arg Pro
    225                 230                 235

GGA GTC TAT ACC AAC ATC AGT CAT CAC TAC AAC TGG ATC CAG TCA ACC       766
Gly Val Tyr Thr Asn Ile Ser His His Tyr Asn Trp Ile Gln Ser Thr
240                 245                 250                 255

ATG ATC CGC AAT GGG CTG CTC AGG CCT GAC CCA GTC CCC TTG CTA CTG       814
Met Ile Arg Asn Gly Leu Leu Arg Pro Asp Pro Val Pro Leu Leu Leu
                260                 265                 270

TTT CTT ACT CTG GCC TGG GCT TCC TCT TTG CTG AGG CCT GCC               856
Phe Leu Thr Leu Ala Trp Ala Ser Ser Leu Leu Arg Pro Ala
            275                 280                 285

TGAGCCCACA CGTGTACGTC ACACCTGTGA GGTCAGGGTG TGTCTCTTTT GTATCTTGCT     916

TGCTAATAAA CCTGTTAATA TTTAAAAAAA AAAAAAAAAA AAA                       959
```

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 285 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

```
Asp Leu Leu Ser Gly Pro Cys Gly His Arg Thr Ile Pro Ser Arg Ile
  1               5                  10                  15

Val Gly Gly Asp Asp Ala Glu Leu Gly Arg Trp Pro Trp Gln Gly Ser
            20                  25                  30

Leu Arg Val Trp Gly Asn His Leu Cys Gly Ala Thr Leu Leu Asn Arg
        35                  40                  45
```

```
Arg Trp Val Leu Thr Ala Ala His Cys Phe Gln Lys Asp Asn Asp Pro
     50                  55                  60

Phe Asp Trp Thr Val Gln Phe Gly Glu Leu Thr Ser Arg Pro Ser Leu
 65                  70                  75                  80

Trp Asn Leu Gln Ala Tyr Ser Asn Arg Tyr Gln Ile Glu Asp Ile Phe
                 85                  90                  95

Leu Ser Pro Lys Tyr Ser Glu Gln Tyr Pro Asn Asp Ile Ala Leu Leu
            100                 105                 110

Lys Leu Ser Ser Pro Val Thr Tyr Asn Asn Phe Ile Gln Pro Ile Cys
        115                 120                 125

Leu Leu Asn Ser Thr Tyr Lys Phe Glu Asn Arg Thr Asp Cys Trp Val
    130                 135                 140

Thr Gly Trp Gly Ala Ile Gly Glu Asp Glu Ser Leu Pro Ser Pro Asn
145                 150                 155                 160

Thr Leu Gln Glu Val Gln Val Ala Ile Ile Asn Asn Ser Met Cys Asn
                165                 170                 175

His Met Tyr Lys Lys Pro Asp Phe Arg Thr Asn Ile Trp Gly Asp Met
            180                 185                 190

Val Cys Ala Gly Thr Pro Glu Gly Gly Lys Asp Ala Cys Phe Gly Asp
        195                 200                 205

Ser Gly Gly Pro Leu Ala Cys Asp Gln Asp Thr Val Trp Tyr Gln Val
    210                 215                 220

Gly Val Val Ser Trp Gly Ile Gly Cys Gly Arg Pro Asn Arg Pro Gly
225                 230                 235                 240

Val Tyr Thr Asn Ile Ser His His Tyr Asn Trp Ile Gln Ser Thr Met
                245                 250                 255

Ile Arg Asn Gly Leu Leu Arg Pro Asp Pro Val Pro Leu Leu Leu Phe
            260                 265                 270
```

Leu Thr Leu Ala Trp Ala Ser Ser Leu Leu Arg Pro Ala
275 280 285

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3866 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

AGTGAGTCTC CTGCCTCAGC CTCCCAAGTA GCTGGGACTT CAGGTGTGTG CCACCATCCT 60

CAGCTAATTT TTTTTTTTTT TTTTTTTTTG AGAAGGAGTC TTGCTCTGTC GCCCAGGCTG 120

GAGTGCAGTG GCGCGATCTT CCAGGCCCCA CCGGGCCCTC AGGAAGGCCT TGCCTACCTG 180

CTTTAAGGGG ACTCCTGGCT CAGGGCCAGG CCCCTGGTGC TGGAGGAGGT GGTGGGTGGA 240

GGGCAGGGGG CACCAAGCGG GCAGCCAGGA CCCCCGGGCT GCAGACAAGA AAAGGACTGT 300

GGGGTCCACC GGGTCTGGGC CACATCAAGG AATGTGGTTG AAGACCCGCC CTTAGGAGCT 360

GAAAGCCAGG GCGCTACCAG GCCTGAGAGG CCCCAAACAG CCCTTGGGCC TGGTTTGGGA 420

GGATTAAGCT GGAGCTCCCA ACCCGCCCTG CCCCCAGGGG GCGACCCCGG GCCCGGCGCG 480

AGAGGAGGCA GAGGGGGCGT CAGGCCGCGG GAGAGGAGGC CATGGGCGCG CGCGGGGCGC 540

TGCTGCTGGC GCTGCTGCTG GCTCGGGCTG GACTCAGGAA GCCGGGTGAG CTCGGGGCGC 600

TGCTGGCGGG ATGGGGAGGC GGGGGAGCGG TGGGGAGGAC GGGAGGTGGA GGCCGCGGGG 660

AGTCACTTCT TGTCTCCCGC AGAGTCGCAG GAGGCGGCGC CGTTATCAGG TAGGGCGCCC 720

AGGACGCGCG ATTCCTGCCA GGGCCGTTGG GCCGAGGTGG ACGGGGGGCG GTGAGGGGGT 780

AGAGGGGGGC CTTTACTGCT CTCTCGCCCC CGCCCCCGGG ATCGAGAACT CTGTTGGCGT 840

GGAAAGTAAC TAACGGACGC TGGAGGGGGA TGGGCGGGCC CTGCAGAGCA CGTGGGAGGA 900

TCTCCAGTGT CACCTACTTC CTGCTGCACA CACGCGAGGG GACCCTGGGT GGGCAAAAAC 960

GTGCTTTCCC GGACGGGGTT GAAGGGGAGA AAGGGAGAGG TCGGGCTTGG GGGGCTGCCT 1020

CCCGCGGCTC AGCAGTTCCT CTGACCATCC GAGGACCATG CGGCCGACGG GTCATCACGT 1080

CGCGCATCGT GGGTGGAGAG GACGCCGAAC TCGGGCGTTG GCCGTGGCAG GGGAGCCTGC 1140

GCCTGTGGGA TTCCCACGTA TGCGGAGTGA GCCTGCTCAG CCACCGCTGG GCACTCACGG 1200

CGGCGCACTG CTTTGAAACG TGAGTGGGGG TGCGAACGGA GGGGTGCGGG GACGGGCAGG 1260

AACAGGGCTG GAGGGAGTGC CACCGAACTT TACCTCTGGT CTGATGCCAG ACTTGGGCGT 1320

GAAAGTTGTG CGTGGATGCG GCCTGGTGTT CTCCTGAGCC CCAGGCTGTG CTGCAGCCGG 1380

TTACACCCAC TCCAGTTCCC TTTGGGTCTC CTGGAGGGAA CCCTGTTCAG GTTATTCCAG 1440

AATGTTCTTC CAGAACATTT CCACACACTT TTGGGTATTC TCTCCCTTTT TCTTTCAACC 1500

CAAAGTTCAC CACTGACCAT CCCACCCTCA TCCCCCCTCC TGGTGGACGG TGCGGTACAG 1560

TGTGGGGCAC TGAGCCAAGG CCAGCACCCC CGGGCCGCTG TGTGGACTCC ATCCTGCCAA 1620

TCCCACATTG GCGTGGTGCA TCTCCCCATT CCTCCTTGGG CTGCATGGGG GTGCCCCTGG 1680

AGGCCTTGGC TCAATGCAAG GCTCCTTGGG ACAGCTCTGG GAGGTGACAA GACCCCACCC 1740

TTCTGCTGCA GGAGCAGGTC CTAGGACTTT GGTTGTGGTC TGTCTGGGCT CCTTCATTTC 1800

TGCAGGGGAC CCTGGGTGTT AGCAAGTAGC AGCAACACCA CAGTTTCCCC TCCTGCACTG 1860

GACCCCAGTT GTGCTCAGGT AGCCAGCCCT CCATCCAGGG CCCCTGACTG CTCTCTTCTC 1920

TTCTGCCAGC TATAGTGACC TTAGTGATCC CTCCGGGTGG ATGGTCCAGT TTGGCCAGCT 1980

GACTTCCATG CCATCCTTCT GGAGCCTGCA GGCCTACTAC ACCCGTTACT TCGTATCGAA 2040

TATCTATCTG AGCCCTCGCT ACCTGGGGAA TTCACCCTAT GACATTGCCT TGGTGAAGCT 2100

GTCTGCACCT GTCACCTACA CTAAACACAT CCAGCCCATC TGTCTCCAGG CCTCCACATT 2160

TGAGTTTGAG AACCGGACAG ACTGCTGGGT GACTGGCTGG GGGTACATCA AAGAGGATGA 2220

GGGTGAGGCT GGGGACAGGC GGGTCAGGGA GGAACTGTCT TTGTTCACCT GTTCCCCTGC 2280

ATAGGCACAA TAGCCCCCTG CTTGGTCTGG GGGTGCAGGC TATGCCCCTC TTGCTTGCAG 2340

TCTCTCCTCA CCTGCCAGGG CAGGGACCAA ACACCCAGTT CTCTCCCTTC CAGGGGCTGT 2400

GGGGGCCAGA AGGAGAGTGT GAGAGGGAGG CCAGTTTGGC GCAAGCCTGT GGGTGGTGCG 2460

GTGGTGGAGG GGTTCTGGAG GGCTTGGCGA CATAAACCTC ATACTTGGAT TTATTCCTGC 2520

ATCTTTCCAC CTCCCCCAGT GCTCACCAAT GCCCCAGGCA TCACCAGGTT GCCCCTTCCC 2580

CCAAGGTCTG GCTTTGGATG CTTATGTGAA CACCGTTTTA AGTTGCCTTG GCCCCTTCCT 2640

CGGTTCCTTT TTGGCTGAGG AATCTCTCCA TGGCTGCAGG CAGGGCCATT GTTGCCATTC 2700

TACAGATAGG GAAAGTGCGG CTGGGGGAGC TCTGACAGCT GTCCCTCCCC GGGGCCTTCT 2760

GTGATGCTGC TGAGGGCCTC TGTTGTGCTG GGGTCTGGGT TGGAGCTGGG GGTAATGGAG 2820

ATGAACCTGC CAGGCACAGT GGGTGCCCCA GGGCCCCCAC CCCCGCAGCC TATGCCATCC 2880

CTCCATAGAG GGGCCTCAGG TTGCTGTCTC TCTCCTTCCC ACTATCGTCC GCACAGCACT 2940

GCCATCTCCC CACACCCTCC AGGAAGTTCA GGTCGCCATC ATAAACAACT CTATGTGCAA 3000

CCACCTCTTC CTCAAGTACA GTTTCCGCAA GGACATCTTT GGAGACATGG TTTGTGCTGG 3060

CAATGCCCAA GGCGGGAAGG ATGCCTGCTT CGTGAGTGTC CTTGCCACCA CTCCCAGCCC 3120

AGGAAAGCAT CCTGTGTCCC TGTGCCTTAT TTGACCCTCA TGCCAACCCC GGGAGGTGGA 3180

GACTGTTGCC CCACTCTGCA GATGCAGAAA CGGAGGCTTG GCTGCTGCCA GGGGGAGGAG 3240

GAGGATGTGC ACCCAGTCTA CCCAGCCCCA TAGCCCTTCC CACTCTCAGC CCCTCCCCTG 3300

CCCCACTCAC TCTGCCCCAG GCTGACCTCA GCCCCGCTGC TCCCCAGGGT GACTCAGGTG 3360

GACCCTTGGC CTGTAACAAG AATGGACTGT GGTATCAGAT TGGAGTCGTG AGCTGGGGAG 3420

TGGGCTGTGG TCGGCCCAAT CGGCCCGGTG TCTACACCAA TATCAGCCAC CACTTTGAGT 3480

GGATCCAGAA GCTGATGGCC CAGAGTGGCA TGTCCCAGCC AGACCCCTCC TGGCCGCTAC 3540

TCTTTTTCCC TCTTCTCTGG GCTCTCCCAC TCCTGGGGCC GGTCTGAGCC TACCTGAGCC 3600

CATGCAGCCT GGGGCCACTG CCAAGTCAGG CCCTGGTTCT CTTCTGTCTT GTTTGGTAAT 3660

AAACACATTC CAGTTGATGC CTTGCAGGGC ATTCTTCAAA AGCAGTGGCT TCATGGACAG 3720

CTCATTCTCT CTTGTGCAGA CAGCCTGTCT GTGCCCCTGG CTCACACCCA CATCTGTTCT 3780

GCACCATAGA ACCATCTGGT TATTTCGATC AGAAAGAGAA TTGTGTGTTG CCCAGGCTGG 3840

TCTTGAACGC CTAGGGTGTC TCGATC 3866

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1165 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

CTGAACCGGG TTGTGGGCGG CGAGGACAGC ACTGACAGCG AGTGGCCCTG GATCGTGAGC 60

ATCCAGAAGA ATGGGACCCA CCACTGCGCA GGTTCTCTGC TCACCAGCCG CTGGGTGATC 120

ACTGCTGCCC ACTGTTTCAA GGACAACCTG AACAAACCAT ACCTGTTCTC TGTGCTGCTG 180

GGGGCCTGGC AGCTGGGGAA CCCTGGCTCT CGGTCCCAGA AGGTGGGTGT TGCCTGGGTG 240

GAGCCCCACC CTGTGTATTC CTGGAAGGAA GGTGCCTGTG CAGACATTGC CCTGGTGCGT 300

CTCGAGCGCT CCATACAGTT CTCAGAGCGG GTCCTGCCCA TCTGCCTACC TGATGCCTCT 360

ATCCACCTCC CTCCAAACAC CCACTGCTGG ATCTCAGGCT GGGGGAGCAT CCAAGATGGA 420

GTTCCCTTGC CCCACCCTCA GACCCTGCAG AAGCTGAAGG TTCCTATCAT CGACTCGGAA 480

GTCTGCAGCC ATCTGTACTG GCGGGGAGCA GGACAGGGAC CCATCACTGA GGACATGCTG 540

TGTGCCGGCT ACTTGGAGGG GGAGCGGGAT GCTTGTCTGG GCGACTCCGG GGGCCCCCTC 600

ATGTGCCAGG TGGACGGCGC CTGGCTGCTG GCCGGCATCA TCAGCTGGGG CGAGGGCTGT 660

GCCGAGCGCA ACAGGCCCGG GGTCTACATC AGCCTCTCTG CGCACCGCTC CTGGGTGGAG 720

AAGATCGTGC AAGGGGTGCA GCTCCGCGGG CGCGCTCAGG GGGGTGGGGC CCTCAGGGCA 780

CCGAGCCAGG GCTCTGGGGC CGCCGCGCGC TCCTAGGGCG CAGCGGGACG CGGGGCTCGG 840

ATCTGAAAGG CGGCCAGATC CACATCTGGA TCTGGATCTG CGGCGGCCTC GGGCGGTTTC 900

CCCCGCCGTA AATAGGCTCA TCTACCTCTA CCTCTGGGGG CCCGGACGGC TGCTGCGGAA 960

AGGAAACCCC CTCCCCGACC CGCCCGACGG CCTCAGGCCC CGCCTCCAAG GCATCAGGCC 1020

CCGCCCAACG GCCTCATGTC CCCGCCCCCA CGACTTCCGG CCCCGCCCCG GGCCCCAGCG 1080

CTTTTGTGTA TATAAATGTT AATGATTTTT ATAGGTATTT GTAACCCTGC CCACATATCT 1140

TATTTATTCC TCCAATTTCA ATAAA 1165

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 933 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

AATGCGGCCA CTCCAAGGAG GCCGGGAGGA TTGTGGGAGG CCAAGACACC CAGGAAGGAC 60

GCTGGCCGTG GCAGGTTGGC CTGTGGTTGA CCTCAGTGGG GCATGTATGT GGGGGCTCCC 120

TCATCCACCC ACGCTGGGTG CTCACAGCCG CCCACTGCTT CCTGAGGTCT GAGGATCCCG 180

GGCTCTACCA TGTTAAAGTC GGAGGGCTGA CACCCTCACT TTCAGAGCCC CACTCGGCCT 240

TGGTGGCTGT GAGGAGGCTC CTGGTCCACT CCTCATACCA TGGGACCACC ACCAGCGGGG 300

ACATTGCCCT GATGGAGCTG GACTCCCCCT TGCAGGCCTC CCAGTTCAGC CCCATCTGCC 360

TCCCAGGACC CCAGACCCCC CTCGCCATTG GGACCGTGTG CTGGGTAAAC GGGCTGGGGG 420

TCCACTCAGG AGAGGCCCTG GCGAGTGTCC TTCAGGAGGT GGCTGTGCCC CTCCTGGACT 480

CGAACATGTG TGAGCTGATG TACCACCTAG GAGAGCCCAG CCTGGCTGGC CAGCGCCTCA 540

TCCAGGACGA CATGCTCTGT GCTGGCTCTG TCCAGGGCAA GAAAGACTCC TGCCAGGGTG 600

ACTCCGGGGG GCCGCTGGTC TGCCCCATCA ATGATACGTG GATCCAGGCC GGCATTGTGA 660

GCTGGGGATT CGGCTGTGCC CGGCCTTTCC GGCCTGGTGT CTACACCCAG GTGCTAAGCT 720

ACACAGACTG GATTCAGAGA ACCCTGGCTG AATCTCACTC AGGCATGTCT GGGGCCCGCC 780

CAGGTGCCCC AGGATCCCAC TCAGGCACCT CCAGATCCCA CCCAGTGCTG CTGCTTGAGC 840

TGTTGACCGT ATGCTTGCTT GGGTCCCTGT GAACCATGAG CCATGGAGTC CGGGATCCCC 900

TTTCTGGTAG GATTGATGGA ATCTAATAAT AAA 933

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 980 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

CCTGTGGTCG CCCCAGGATG CTGAACCGAA TGGTGGGCGG GCAGGACACG CAGGAGGGCG 60

AGTGGCCCTG GCAAGTCAGC ATCCAGCGCA ACGGAAGCCA CTTCTGCGGG GGCAGCCTCA 120

TCGCGGAGCA GTGGGTCCTG ACGGCTGCGC ACTGCTTCCG CAACACCTCT GAGACGTCCC 180

TGTACCAGGT CCTGCTGGGG GCAAGGCAGC TAGTGCAGCC GGGACCACAC GCTATGTATG 240

CCCGGGTGAG GCAGGTGGAG AGCAACCCCC TGTACCAGGG CACGGCCTCC AGCGCTGACG 300

TGGCCCTGGT GGAGCTGGAG GCACCAGTGC CCTTCACCAA TTACATCCTC CCCGTGTGCC 360

TGCCTGACCC CTCGGTGATC TTTGAGACGG GCATGAACTG CTGGGTCACT GGCTGGGGCA 420

GCCCCAGTGA GGAAGACCTC CTGCCCGAAC CGCGGATCCT GCAGAAACTC GCTGTGCCCA 480

TCATCGACAC ACCCAAGTGC AACCTGCTCT ACAGCAAAGA CACCGAGTTT GGCTACCAAC 540

CCAAAACCAT CAAGAATGAC ATGCTGTGCG CCGGCTTCGA GGAGGGCAAG AAGGATGCCT 600

GCAAGGGCGA CTCGGGCGGC CCCCTGGTGT GCCTCGTGGG TCAGTCGTGG CTGCAGGCGG 660

GGGTGATCAG CTGGGGTGAG GGCTGTGCCC GCCAGAACCG CCCAGGTGTC TACATCCGTG 720

TCACCGCCCA CCACAACTGG ATCCATCGGA TCATCCCCAA ACTGCAGTTC CAGCCAGCGA 780

GGTTGGGCGG CCAGAAGTGA GACCCCCGGG GCCAGGAGCC CCTTGAGCAG AGCTCTGCAC 840

CCAGCCTGCC CGCCCACACC ATCCTGCTGG TCCTCCCAGC GCTGCTGTTG CACCTGTGAG 900

CCCCACCAGA CTCATTTGTA AATAGCGCTC CTTCCTCCCC TCTCAAATAC CCTTATTTTA 960

TTTATGTTTC TCCCAATAAA 980